



OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/084,018

DATE: 05/07/2002
TIME: 11:37:44

Input Set : N:\Crf3\RULE60\10084018.raw
Output Set: N:\CRF3\05072002\J084018.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Bandman, Olga
6 Hawkins, Phillip R.
7 Hillman, Jennifer L.
8 Lal, Preeti
9 Goli, Surya K.

11 (ii) TITLE OF INVENTION: NOVEL HUMAN SERINE
12 CARBOXYPEPTIDASE

14 (iii) NUMBER OF SEQUENCES: 8

16 (iv) CORRESPONDENCE ADDRESS:
17 (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
18 (B) STREET: 3174 Porter Drive
19 (C) CITY: Palo Alto
20 (D) STATE: CA
21 (E) COUNTRY: USA
22 (F) ZIP: 94304

24 (v) COMPUTER READABLE FORM:
25 (A) MEDIUM TYPE: Diskette
26 (B) COMPUTER: IBM Compatible
27 (C) OPERATING SYSTEM: DOS
28 (D) SOFTWARE: FastSEQ for Windows Version 2.0

30 (vi) CURRENT APPLICATION DATA:
C--> 31 (A) APPLICATION NUMBER: US/10/084,018
C--> 32 (B) FILING DATE: 25-Feb-2002

34 (vii) PRIOR APPLICATION DATA:
36 (A) APPLICATION NUMBER: US/09/299,689A
37 (B) FILING DATE:
39 (A) APPLICATION NUMBER: 08/828,488
40 (B) FILING DATE:

43 (viii) ATTORNEY/AGENT INFORMATION:
44 (A) NAME: Billings, Lucy J.
45 (B) REGISTRATION NUMBER: 36,749
46 (C) REFERENCE/DOCKET NUMBER: PF-0241 US

48 (ix) TELECOMMUNICATION INFORMATION:
49 (A) TELEPHONE: 415-855 0555
50 (B) TELEFAX: 415-845-4166

53 (2) INFORMATION FOR SEQ ID NO: 1:
55 (i) SEQUENCE CHARACTERISTICS:
56 (A) LENGTH: 477 amino acids
57 (B) TYPE: amino acid
58 (C) STRANDEDNESS: single
59 (D) TOPOLOGY: linear

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/084,018

DATE: 05/07/2002

TIME: 11:37:44

Input Set : N:\Crf3\RULE60\10084018.raw
 Output Set: N:\CRF3\05072002\J084018.raw

61 (vii) IMMEDIATE SOURCE:
 62 (A) LIBRARY: MPHGN0T03
 63 (B) CLONE: 443004
 65 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 67 Met Val Gly Ala Met Trp Lys Val Ile Val Ser Leu Val Leu Met
 68 1 5 10 15
 69 Pro Gly Pro Cys Asp Gly Leu Phe His Ser Leu Tyr Arg Ser Val Ser
 70 20 25 30
 71 Met Pro Pro Lys Gly Asp Ser Gly Gln Pro Leu Phe Leu Thr Pro Tyr
 72 35 40 45
 73 Ile Glu Ala Gly Lys Ile Gln Lys Gly Arg Glu Leu Ser Leu Val Gly
 74 50 55 60
 75 Pro Phe Pro Gly Leu Asn Met Lys Ser Tyr Ala Asp Phe Leu Thr Val
 76 65 70 75 80
 77 Asn Lys Thr Tyr Asn Ser Asn Leu Phe Phe Trp Phe Pro Ala Gln
 78 85 90 95
 79 Ile Gln Pro Glu Asp Ala Pro Val Val Leu Trp Leu Gln Gly Glu Pro
 80 100 105 110
 81 Gly Gly Ser Ser Met Phe Gly Leu Phe Val Glu His Gly Pro Tyr Val
 82 115 120 125
 83 Val Thr Ser Asn Met Thr Leu Arg Asp Arg Asp Phe Pro Trp Thr Thr
 84 130 135 140
 85 Thr Leu Ser Met Leu Tyr Ile Asp Asn Pro Val Gly Thr Gly Phe Ser
 86 145 150 155 160
 87 Phe Thr Asp Asp Thr His Gly Tyr Ala Val Asn Glu Asp Asp Val Ala
 88 165 170 175
 89 Arg Asp Leu Tyr Ser Ala Leu Ile Gln Phe Phe Gln Ile Phe Pro Glu
 90 180 185 190
 91 Tyr Lys Asn Asn Asp Phe Tyr Val Thr Gly Glu Ser Tyr Ala Gly Lys
 92 195 200 205
 93 Tyr Val Pro Ala Ile Ala His Leu Ile His Ser Leu Asn Pro Val Arg
 94 210 215 220
 95 Glu Val Lys Ile Asn Leu Asn Gly Ile Ala Ile Gly Asp Gly Tyr Ser
 96 225 230 235 240
 97 Asp Pro Glu Ser Ile Ile Gly Gly Tyr Ala Glu Phe Leu Tyr Gln Ile
 98 245 250 255
 99 Gly Leu Leu Asp Glu Lys Gln Lys Lys Tyr Phe Gln Lys Gln Cys His
 100 260 265 270
 101 Glu Cys Ile Glu His Ile Arg Lys Gln Asn Trp Phe Glu Ala Phe Glu
 102 275 280 285
 103 Ile Leu Asp Lys Leu Leu Asp Gly Asp Leu Thr Ser Asp Pro Ser Tyr
 104 290 295 300
 105 Phe Gln Asn Val Thr Gly Cys Ser Asn Tyr Tyr Asn Phe Leu Arg Cys
 106 305 310 315 320
 107 Thr Glu Pro Glu Asp Gln Leu Tyr Tyr Val Lys Phe Leu Ser Leu Pro
 108 325 330 335
 109 Glu Val Arg Gln Ala Ile His Val Gly Asn Gln Thr Phe Asn Asp Gly
 110 340 345 350
 111 Thr Ile Val Glu Lys Tyr Leu Arg Glu Asp Thr Val Gln Ser Val Lys

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/084,018

DATE: 05/07/2002
TIME: 11:37:44

Input Set : N:\CrF3\RULE60\10084018.raw
Output Set: N:\CRF3\05072002\J084018.raw

112 355 360 365
113 Pro Trp Leu Thr Glu Ile Met Asn Asn Tyr Lys Val Leu Ile Tyr Asn
114 370 375 380
115 Gly Gln Leu Asp Ile Ile Val Ala Ala Ala Leu Thr Glu Arg Ser Leu
116 385 390 395 400
117 Met Gly Met Asp Trp Lys Gly Ser Gln Glu Tyr Lys Lys Ala Glu Lys
118 405 410 415
119 Lys Lys Val Trp Lys Ile Phe Lys Ser Asp Ser Gly Val Ala Gly Tyr
120 420 425 430
121 Ile Arg Gln Val Gly Asp Phe His Gln Val Ile Ile Arg Gly Gly Gly
122 435 440 445
123 His Thr Leu Pro Tyr Asp Gln Pro Leu Arg Ala Phe Asp Met Ile Asn
124 450 455 460
125 Arg Phe Ile Tyr Gly Lys Gly Trp Asp Pro Tyr Val Gly
126 465 470 475

128 (2) INFORMATION FOR SEQ ID NO: 2:

130 (i) SEQUENCE CHARACTERISTICS:
131 (A) LENGTH: 1670 base pairs
132 (B) TYPE: nucleic acid
133 (C) STRANDEDNESS: single
134 (D) TOPOLOGY: linear
136 (vii) IMMEDIATE SOURCE:
137 (A) LIBRARY: MPHGN0T03
138 (B) CLONE: 443004

140 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
142 AAGCGCTGCA AGGACAACCG GCTGGGTCC TTGCGCGCCG GGCTCAGGGA GGAGCACCAGA 60
143 CTGCGCCGCA CCCTGAGAGA TGGTTGGTGC CATGTGGAAG GTGATTGTT CGCTGGTCCT 120
144 GTTGATGCCT GGCCCCCTGTG ATGGGCTGTT TCACCTCCCTA TACAGAACTG TTTCCATGCC 180
145 ACCTAAGGGGA GACTCAGGAC AGCCATTATT TCTCACCCCCCT TACATTGAAG CTGGGAAGAT 240
146 CCAAAAAGGA AGAGAATTGA GTTTGGTCGG TCCTTTCCCA GGACTGAACA TGAAGAGTTA 300
147 TGGCGACTTC CTCACTGTGA ATAAGACTTA CAACAGCAAC CTCTTCTTCTI GGTCTTCCC 360
148 AGCICAGATA CAGCCAGAAG ATGCCCCAGT AGTCTCTGG CTACAGGGTG AGCCGGGAGG 420
149 TTCACTCATG TTGGACTCTI TIGTGGAAACA TGGCCCTTAT GTGICACAA GAAACATGAC 480
150 CTTGCGTGAC AGAGACTTCC CCTGGACCCAC AACGCTCTCC ATGCTTACA TTGACAATCC 540
151 AGTGGGCACA GGCTTCAGTT TIACTGTGA TACCCACGGGA TATGCAGTCA ATGAGGACGA 600
152 TGTAGCACCG GATTTATACA GIGCACTAAT TCAGTTTTTC CAGATATTTC CTGAATATAA 660
153 AAATAATGAC TTTTATGTCA CTGGGGAGTC TTATGCAAGGG AAATATGTGC CAGCCATTGC 720
154 ACACCTCATC CATTCCCTCA ACCCTGTGAG AGAGGTGAAG ATCAACCTGA ACGGAATTGC 780
155 TATTGGAGAT GGATATTCTG ATCCCGAATC AATTATAGGG GGCTATGCAAG AATTCCCTGTA 840
156 CCAAATTGGC TTGTTGGATG AGAAGCAAA AAAGTACTTC CAGAAGCAGT GCCATGAATG 900
157 CATAGAACAC ATCAGGAAGC AGAACTGTT TGAGGCCTTT GAAATACTGG ATAAACTACT 960
158 AGATGGCGAC TTAACAAAGTG ATCCTTCTTA CTTCCAGAAT GTTACAGGAT GTAGTAATTA 1020
159 CTATAACTTT TTGCGGTGCA CGGAACCTGA GGATCAGCTT TACTATGTGA AATTTTGTC 1080
160 ACTCCCAGAG GTGAGACAAG CCATCCACGT GGGGAATCAG ACTTTIAATG ATGGAACAT 1140
161 AGTTGAAAAG TACTTGCGAG AAGATAACAGT ACAGTCAGTT AAGCCATGGT TAATCTGAAT 1200
162 CATGAATAAT TATAAGGTTC TGATCTACAA TGGCCAATG GACATCATCG TGGCAGCTGC 1260
163 CCTGACAGAG CGCTCCTTGA TGGCATGGA CTGAAAGGA TCCCAAGGAAT ACAAGAAGGC 1320
164 AGAAAAAAAGA AAAGTTTGGGA AGATCTTAA ATCTGACAGT GGAGTGGCTG GTTACATCCG 1380
165 GCAAGTGGGT GACTTCCATC AGGTAATTAT TCGAGGTGGA GGACATACTT TACCCATG 1440

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/084,018

DATE: 05/07/2002

TIME: 11:37:44

Input Set : N:\Crf3\RULE60\10084018.raw
 Output Set: N:\CRF3\05072002\J084018.raw

166 CCAGCCTCTG AGAGCTTTG ACATGATTAA TCGATTCAATTATGGAAAAG GATGGGATCC 1500
 167 TTATGTTGGA TAAACTACCT TCCCAAAAGA GAACATCAGA GGTTTCATT GCTGAAAAGA 1560
 168 AAATCGTAAA AACAGAAAAT GTCATAGGAA TAAAAAAATT ATCTTTCAT ATCTGCAAGA 1620
 169 TCTTITTCAT CAATAAAAAT TATCCTTGAA ACAAAAAAAA AAAGAAAAAG 1670
 171 (2) INFORMATION FOR SEQ ID NO: 3:
 173 (i) SEQUENCE CHARACTERISTICS:
 174 (A) LENGTH: 476 amino acids
 175 (B) TYPE: amino acid
 176 (C) STRANDEDNESS: single
 177 (D) TOPOLOGY: linear
 179 (vii) IMMEDIATE SOURCE:
 180 (A) LIBRARY: MMLR3DT01
 181 (B) CLONE: 566993
 183 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
 185 Met Val Gly Ala Met Trp Lys Val Ile Val Ser Leu Val Leu Leu Met
 186 1 5 10 15
 187 Pro Gly Pro Cys Asp Gly Leu Phe His Ser Leu Tyr Arg Ser Val Ser
 188 20 25 30
 189 Met Pro Pro Lys Gly Asp Ser Gly Gln Pro Leu Phe Leu Thr Pro Tyr
 190 35 40 45
 191 Ile Glu Ala Gly Lys Ile Gln Lys Gly Arg Glu Leu Ser Leu Val Gly
 192 50 55 60
 193 Pro Phe Pro Gly Leu Asn Met Lys Ser Tyr Ala Gly Phe Leu Thr Val
 194 65 70 75 80
 195 Asn Lys Thr Tyr Asn Ser Asn Leu Phe Phe Trp Phe Pro Ala Gln
 196 85 90 95
 197 Ile Gln Pro Glu Asp Ala Pro Val Val Leu Trp Leu Gln Gly Gly Pro
 198 100 105 110
 W--> 199 Gly Gly Ser Ser Met Xaa Gly Leu Phe Val Glu His Gly Pro Tyr Val
 200 115 120 125
 201 Val Thr Ser Asn Met Thr Leu Arg Asp Arg Asp Phe Pro Trp Thr Thr
 202 130 135 140
 W--> 203 Thr Xaa Ser Met Leu Tyr Ile Asp Asn Pro Val Gly Thr Gly Phe Ser
 204 145 150 155 160
 205 Phe Thr Asp Asp Thr His Gly Tyr Ala Val Asn Glu Asp Asp Val Ala
 206 165 170 175
 207 Arg Asp Leu Tyr Ser Ala Leu Ile Gln Phe Phe Gln Ile Phe Pro Glu
 208 180 185 190
 209 Tyr Lys Asn Asn Asp Phe Tyr Val Thr Gly Glu Ser Tyr Ala Gly Lys
 210 195 200 205
 211 Tyr Val Pro Ala Ile Ala His Leu Ile His Ser Leu Asn Pro Val Arg
 212 210 215 220
 213 Glu Val Lys Ile Asn Leu Asn Gly Ile Ala Ile Gly Asp Gly Tyr Ser
 214 225 230 235 240
 215 Asp Pro Glu Ser Ile Ile Gly Gly Tyr Ala Glu Phe Leu Tyr Gln Ile
 216 245 250 255
 217 Gly Leu Leu Asp Glu Lys Gln Lys Lys Tyr Phe Gln Lys Gln Cys His
 218 260 265 270
 219 Glu Cys Ile Glu His Ile Arg Lys Gln Asn Trp Phe Glu Ala Phe Glu

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/084,018

DATE: 05/07/2002
TIME: 11:37:44

Input Set : N:\CrF3\RULE60\10084018.raw
Output Set: N:\CRF3\05072002\J084018.raw

220	275	280	285
221	Ile Leu Asp Lys Leu Leu Asp Gly Asp Leu Thr Ser Asp Pro Ser Tyr		
222	290	295	300
223	Phe Gln Asn Val Thr Gly Cys Ser Asn Tyr Tyr Asn Phe Leu Arg Cys		
224	305	310	315
225	Thr Glu Pro Glu Asp Gln Leu Tyr Tyr Val Lys Phe Leu Ser Leu Pro		320
226	325	330	335
227	Glu Val Arg Gln Ala Ile His Val Gly Asn Gln Thr Phe Asn Asp Gly		
228	340	345	350
229	Thr Ile Val Glu Lys Tyr Leu Arg Glu Asp Thr Val Gln Ser Val Lys		
230	355	360	365
231	Pro Trp Leu Thr Glu Ile Met Asn Asn Tyr Lys Val Leu Ile Tyr Asn		
232	370	375	380
233	Gly Gln Leu Asp Ile Ile Val Ala Ala Ala Leu Thr Glu Arg Ser Leu		
234	385	390	395
235	Met Gly Met Asp Trp Lys Gly Ser Gln Glu Tyr Lys Lys Ala Glu Lys		400
236	405	410	415
237	Lys Val Trp Lys Ile Phe Lys Ser Asp Ser Glu Val Ala Gly Tyr Ile		
238	420	425	430
239	Arg Gln Val Gly Asp Phe His Gln Val Ile Ile Arg Gly Gly Gly His		
240	435	440	445
241	Ile Leu Pro Tyr Asp Gln Pro Leu Arg Ala Phe Asp Met Ile Asn Arg		
242	450	455	460
243	Phe Ile Tyr Gly Lys Gly Trp Asp Pro Tyr Val Gly		
244	465	470	475
246	(2) INFORMATION FOR SEQ ID NO: 4:		
248	(i) SEQUENCE CHARACTERISTICS:		
249	(A) LENGTH: 1551 base pairs		
250	(B) TYPE: nucleic acid		
251	(C) STRANDEDNESS: single		
252	(D) TOPOLOGY: linear		
254	(vii) IMMEDIATE SOURCE:		
255	(A) LIBRARY: MMLR3DT01		
256	(B) CLONE: 566993		
258	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:		
260	GAAAGCTGGT ACGCCTGCNG GTNCCGGTCC GGAATTNCNG GGTNGACCCA CGCGTCCGAN	60	
261	CGACTGCGCC GCACCCCTGAG AGATGGTGG TGCCATGTGG AAGGTGATTG TTTCGCTGGT	120	
262	CCTGTTGATG CCTGGCCCCCT GTGATGGGCT GTTCACTCC CTATACAGAA GTGTTTCCAT	180	
263	GCCACCTAAG GGAGACTCAG GACAGCCATT ATTTCACC CTTTACATTG AAGCTGGGAA	240	
264	GATCCAAAAA GGAAGAGAAAT TGAGTTGGT CGGCCCTTTC CCAGGACTGAA ACATGAAGAG	300	
265	TTATGCCGGC TTCTTCACCG TGAATAAGAC TTACAACAGC AACCTCTTCT TCTGGTTCTT	360	
266	CCCAGCTCAG ATACAGCCAG AAGATGCCCG AGTAGTTCTC TGGCTACAGG GTGGGCCGGG	420	
267	AGGTTCATCC ATGTTWGGAC TCTTTGTGGA ACATGGGCCT TATGTTGTCA CAAGTAACAT	480	
268	GACCTTGCGT GACAGAGACT TCCCCTGGAC CACAACGSTC TCCATGCTT ACATTGACAA	540	
269	TCCAGTGGGC ACAGGCTTCA GTTTTACTGA TGATACCCAC GGATATGCGAG TCAATGAGGA	600	
270	CGATGTAGCA CGGGATTAT ACAGTGCCT AATTCACTT TTCCAGATAT TICCTGAATA	660	
271	TAAAAATAAT GACTTTATG TCACTGGGA GTCTTATGCA GGGAAATATG TGCCAGCCAT	720	
272	TGCACACCTC ATCCATTCCC TCAACCCGT GAGAGAGGTG AAGATCAACC TGAACGGAAT	780	
273	TGCTATTGGA GATGGATATT CTGATCCGA ATCAATTATA GGGGGCTATG CAGAATTCCCT	840	

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 05/07/2002
PATENT APPLICATION: US/10/084,018 TIME: 11:37:45

Input Set : N:\CrF3\RULE60\10084018.raw
Output Set: N:\CRF3\05072002\J084018.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:3; Xaa Pos.118,146
Seq#:4; N Pos. 19,23,38,44,60,1550

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/084,018

DATE: 05/07/2002

TIME: 11:37:45

Input Set : N:\Crf3\RULE60\10084018.raw
Output Set: N:\CRF3\05072002\J084018.raw

L:31 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:32 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]

L:199 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:112

L:203 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:144